

OIPF

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/848,852A

DATE: 12/20/2001

TIME: 11:52:09

Input Set : N:\Crf3\RULE60\09848852.raw

Output Set: N:\CRF3\12202001\I848852A.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Hillman, Jennifer L.

6 Tang, Y. Tom

7 Corley, Neil C.

8 Guegler, Karl J.

9 Yue, Henry

10 Patterson, Chandra

12 (ii) TITLE OF INVENTION: PROTEINS ASSOCIATED WITH EGF SIGNALING

14 (iii) NUMBER OF SEQUENCES: 5

16 (iv) CORRESPONDENCE ADDRESS:

17 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

18 (B) STREET: 3174 Porter Dr.

19 (C) CITY: Palo Alto

20 (D) STATE: CA

21 (E) COUNTRY: USA

22 (F) ZIP: 94304

24 (v) COMPUTER READABLE FORM:

25 (A) MEDIUM TYPE: Diskette

26 (B) COMPUTER: IBM Compatible

27 (C) OPERATING SYSTEM: DOS

28 (D) SOFTWARE: FastSEQ for Windows Version 2.0

30 (vi) CURRENT APPLICATION DATA:

C--> 31 (A) APPLICATION NUMBER: US/09/848,852A

C--> 32 (B) FILING DATE: 07-May-2001

34 (vii) PRIOR APPLICATION DATA:

35 (A) APPLICATION NUMBER: 09/069,725

36 (B) FILING DATE:

40 (viii) ATTORNEY/AGENT INFORMATION:

41 (A) NAME: Billings, Lucy J.

42 (B) REGISTRATION NUMBER: 36,749

43 (C) REFERENCE/DOCKET NUMBER: PF-0515 US

45 (ix) TELECOMMUNICATION INFORMATION:

46 (A) TELEPHONE: 650-855-0555

47 (B) TELEFAX: 650-845-4166

50 (2) INFORMATION FOR SEQ ID NO: 1:

52 (i) SEQUENCE CHARACTERISTICS:

53 (A) LENGTH: 298 amino acids

54 (B) TYPE: amino acid

55 (C) STRANDEDNESS: single

56 (D) TOPOLOGY: linear

58 (vii) IMMEDIATE SOURCE:

59 (A) LIBRARY: COLNNOT16

60 (B) CLONE: 1281694

62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

64 Met Asn Asp Ser Leu Arg Thr Asp Val Phe Val Arg Phe Gln Pro Glu

65 1

5

10

15

RECEIVED

JAN 18 2002

TECH CENTER 1600/2900

ENTERED

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/848,852A

TIME: 11:52:09

Input Set : N:\Crif3\RULE60\09848852.raw

Output Set: N:\CRF3\12202001\I848852A.raw

```

66 Ser Ile Ala Cys Ala Cys Ile Tyr Leu Ala Ala Arg Thr Leu Glu Ile
67          20          25          30
68 Pro Leu Pro Asn Arg Pro His Trp Phe Leu Leu Phe Gly Ala Thr Glu
69          35          40          45
70 Glu Glu Ile Gln Glu Ile Cys Leu Lys Ile Leu Gln Leu Tyr Ala Arg
71          50          55          60
72 Lys Lys Val Asp Leu Thr His Leu Glu Gly Glu Val Glu Lys Arg Lys
73          65          70          75          80
74 His Ala Ile Glu Glu Ala Lys Ala Gln Ala Arg Gly Leu Leu Pro Gly
75          85          90          95
76 Gly Thr Gln Val Leu Asp Gly Thr Ser Gly Phe Ser Pro Ala Pro Lys
77          100         105         110
78 Leu Val Glu Ser Pro Lys Glu Gly Lys Gly Ser Lys Pro Ser Pro Leu
79          115         120         125
80 Ser Val Lys Asn Thr Lys Arg Arg Leu Glu Gly Ala Lys Lys Ala Lys
81          130         135         140
82 Ala Asp Ser Pro Val Asn Gly Leu Pro Lys Gly Arg Glu Ser Arg Ser
83          145         150         155         160
84 Arg Ser Arg Ser Arg Glu Gln Ser Tyr Ser Arg Ser Pro Ser Arg Ser
85          165         170         175
86 Ala Ser Pro Lys Arg Arg Lys Ser Asp Ser Gly Ser Thr Ser Gly Gly
87          180         185         190
88 Ser Lys Ser Gln Ser Arg Ser Arg Ser Arg Ser Asp Ser Pro Pro Arg
89          195         200         205
90 Gln Ala Pro Arg Ser Ala Pro Tyr Lys Gly Ser Glu Ile Arg Gly Ser
91          210         215         220
92 Arg Lys Ser Lys Asp Cys Lys Tyr Pro Gln Lys Pro His Lys Ser Arg
93          225         230         235         240
94 Ser Arg Ser Ser Ser Arg Ser Arg Ser Arg Ser Arg Glu Arg Ala Asp
95          245         250         255
96 Asn Pro Gly Lys Tyr Lys Lys Lys Ser His Tyr Tyr Arg Asp Gln Arg
97          260         265         270
98 Arg Glu Arg Ser Arg Ser Tyr Glu Arg Thr Gly Arg Arg Tyr Glu Arg
99          275         280         285
100 Asp His Pro Gly His Ser Arg His Arg Arg
101          290         295
103 (2) INFORMATION FOR SEQ ID NO: 2:
105   (i) SEQUENCE CHARACTERISTICS:
106       (A) LENGTH: 3464 base pairs
107       (B) TYPE: nucleic acid
108       (C) STRANDEDNESS: single
109       (D) TOPOLOGY: linear
111   (vii) IMMEDIATE SOURCE:
112       (A) LIBRARY: COLNNOT16
113       (B) CLONE: 1281694
115   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
117 TTCGTGAAGC ACTCCATGGA GCATGTGTCA ATGGCCTGTG TCCACCTGGC TTCCAAGATA      60
118 GAAGAGGCC CCAAGACGCAT ACGGGACGTC ATCAATGTGT TTCACCGCCT TCGACAGCTG      120
119 AGAGACAAAA AATAATCGTT ATGTACCTTC AGGTGTTAGA GTGTGAGCGT AACCAACACC      180

```

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/848,852A

TIME: 11:52:09

Input Set : N:\Crif3\RULE60\09848852.raw

Output Set: N:\CRF3\12202001\I848852A.raw

120	TGGTCCAGAC	CTCATGGGTA	GCCTCTGAGG	GTAAGTGA	CTCTGCTGTC	240
121	CAAGCGCTTT	GGTGCAGGGA	CAGCGGCATC	TTCAGCCAAT	CCAGTGCAGG	300
122	AAGGCTGGCT	CTAGACTGGT	GACCCCTTGT	TGAAATGGGA	CAGTTGGCAG	360
123	GAGCCCGAGA	AGAGGCCTGC	CCTTGGGGTGC	GGAGTCTCCC	TCCGCACGAT	420
124	GTCCAACCTG	CACCCAAGGG	GCTTTTCCCT	CTTCCAAGTG	GACTCCTTCA	480
125	AGCTCGGTCA	GCAGAGAAGG	GGCCTGCCGC	CAGCGCCCTG	GAGGAAGAGG	540
126	CAAGAGGATG	GCTTGTCTCC	CAGCAGCCAC	ACCGGCTTTG	TGCTCAGCCA	600
127	GTTTGCATGT	TTCTCTGCAC	TATGGATTTT	GAGCATTTAG	ATTTCTTTAA	660
128	TTTAGTGACT	CCAGTAGACA	TTTTCTTTCT	GAGGCATCGT	GCTTTGCATG	720
129	AAGGTTGAGG	GGAAAAGTAA	AGTTAAAGTC	GGTTCTCTTT	CATAGCAACA	780
130	GACATTCAGC	CAGCTTTTTT	TTTTTCTAAT	AATTTCTGTG	CCTTTCTGTC	840
131	TGTATTTAGA	AAAAGCAGCT	AGAATATTTT	TCCATTAAGT	CTTGAGATTC	900
132	CTAGCTCTGA	GTCCTAGCAA	TAGACTCCTT	ACGAGGAGTAG	TACGTTTATC	960
133	CTAGATAATG	CAGGCGGAAG	ACCTGGGGTTC	ACGGGTGGGG	CATTGCAGTT	1020
134	TTGGCTTCCA	GGAATTACAT	GAACGACAGC	CTTCGCACCG	ACGTCTTCGT	1080
135	CCAGAGAGCA	TCGCCTGTGC	CTGCATTTAT	CTTGCTGCCC	GGACGCTGGA	1140
136	CCCAATCGTC	CCCATTTGGT	TCTTTTGT	GGAGCAACTG	AAGAAGAAAT	1200
137	TGCTTAAAGA	TCTTGCAGCT	TTATGCTCGG	AAAAAGGTTG	ATCTCACACA	1260
138	GAAGTGGA	AAAGAAAGCA	CGCTATCGAA	GAGGCAAAGG	CCCAAGCCCG	1320
139	CCTGGGGGCA	CACAGGTGCT	GGATGGTACC	TCGGGGTTCT	CTCCTGCCCC	1380
140	GAATCCCCCA	AAGAAGGTAA	AGGGAGCAAG	CCTTCCCCAC	TGTCTGTGAA	1440
141	AGGAGGCTGG	AGGGCGCCAA	GAAAGCCAAG	GCGGACAGCC	CCGTGAACGG	1500
142	GGGCGAGAGA	GTCGGAGTCG	GAGCCGGAGC	CGTGAGCAGA	GCTACTCGAG	1560
143	CGATCAGCGT	CTCCTAAGAG	GAGGAAAAGT	GACAGCGGCT	CCACATCTGG	1620
144	TCGCAGAGCC	GCTCCCGGAG	CAGGAGTGAC	TCCCCACCGA	GACAGGCCCC	1680
145	CCCTACAAAG	GCTCTGAGAT	TCGGGGCTCC	CGGAAGTCCA	AGGACTGCAA	1740
146	AAGCCACACA	AGTCTCGGAG	CCGGAGTTCT	TCCCCTTCTC	GAAGCAGGTC	1800
147	GCGGATAATC	CGGGAATAA	CAAGAAGAAA	AGTCATTACT	ACAGAGATCA	1860
148	CGCTCGAGGT	CGTAGAACG	CACAGGCCGT	CGCTATGAGC	GGGACCACCC	1920
149	AGGCATCGGA	GGTGAGCCGG	GGTTGCAGTG	ACTGGTGGCC	GCAAGCCCTT	1980
150	TACCTGATGG	GTCGCCCTTT	ACCCCGGGTG	GCTGCCCTTT	GACCCCGGGG	2040
151	GCGCAAGTGG	TCCTAGAACA	GGATTCTTTT	TGGAAATGTC	TGTCGACTGG	2100
152	ATTTGGAAAT	GGAAGTGAAG	GACCGGTGAC	ACGTGCTTCA	GACCGGTCTG	2160
153	CACACCTGGG	CCCGTGCAGG	GCTCAGCTCG	GCAGCAGCTC	TGAGGGCAGC	2220
154	AGTGAATGCA	CACGCCCTTG	TTGGCGTGGC	CTGGCATGGC	CTGGTGCTAT	2280
155	TCTCCACTCC	CCGACTGATA	CTCAATTACG	TGAAGCCAAG	AAAGATGATT	2340
156	TTGCCTATAT	TAGGTTGTAC	TTATGTACAT	ATTTTGCAGT	GTTTCACAGG	2400
157	CTTAAGTACC	CCTTATTCTC	TCTCCACGTT	GTAAATAAAC	ATGTGTTTAA	2460
158	AGCTATGTAT	GAAAAGTCA	AACTTGAATC	CCGTACAGCT	AAAAGTGTG	2520
159	TGACTTTTAA	AATGTGAGGG	TATTTGGATC	TGTGTTGAAA	GTCGTATATT	2580
160	CGGTGCTGAG	TGCAGGCCAC	CAGCTCCTAA	ATAGAGGTTT	CCTATATGCG	2640
161	GGTGAATAAA	CACAACTCTC	TCCACTCAGG	ACATCCGGAG	CGTTATGGAC	2700
162	GTCGTTCTGT	GTGCTTGTGA	AAGTGTCCAG	GCGTGTGCAC	AGCCAGTGCG	2760
163	GGCTCCTTGC	TCCCTGCTGT	ACTGAAGTTT	TGGATTTTGC	ATCCAATCCT	2820
164	CCTTCTGCCG	AAGCTTGTGA	GGGGCCTGAG	TCCTCTGCCC	ATCAGGATGA	2880
165	CCTGCAGGGG	CATAGGAGG	AAGTTTGGGA	AACACAGAAT	GATTCCAAGG	2940
166	CCTGAGGGGG	ACTGGTTTGT	AACCCATGAC	ATCTGTGGGC	GAGAGAGGCA	3000
167	GGACACTTGG	AGGGTCACCC	CACGGGGGTG	GCACCTGCAC	TCTGAGTGCC	3060
168	ATCAGCTGCC	TCTTACCGTG	GACACAGTTT	TGGTTTTGGG	GACTAGGGGG	3120

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/848,852A

TIME: 11:52:09

Input Set : N:\Crf3\RULE60\09848852.raw

Output Set: N:\CRF3\12202001\I848852A.raw

```

169 GGTGGTACCG TTTGGACTTA CTAGGGCAGT GGGACATATA GGCCGGGGCT AGTGGGATAA 3180
170 CGGGGAGTTA CGCCTGATGA CTTTTTTGAT GGAATCCTGC ATTAGATAGC TGGTGGGACC 3240
171 CCCCCCTCAG AATTGGGGAA CTGAGGAGAC TCCAGGGAGG GTGTCCTTCC AGGGAGAGCA 3300
172 GCTATGAGGG GCCCCCTAGC TTCCTGTGCC TGGGAAGTAAG AGAACCAGTA AAGGGCCATA 3360
173 CACACCTGTA CCAAGAGAC CGCTCTCCAT TTGCTTTCTT TTTTACTAA ATAATTGTAA 3420
174 AATATTATTA TGACATAAAG AACCATTAA GGCCAAAAAA AAAA 3464

```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 332 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: BEPINOT01

(B) CLONE: 2056178

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

190 Met Tyr Ser Ala Gln Arg Phe Trp Gly Thr Ile Trp Ala Arg Arg Gly
191 1 5 10 15
192 Ala His Leu Ala Pro Pro Asp Ala Ser Ile Leu Ile Ser Asn Val Cys
193 20 25 30
194 Ser Ile Gly Asp His Val Ala Gln Glu Leu Phe Gln Gly Ser Asp Leu
195 35 40 45
196 Gly Met Ala Glu Glu Ala Glu Arg Pro Gly Glu Lys Ala Gly Gln His
197 50 55 60
198 Ser Pro Leu Arg Glu Glu His Val Thr Cys Val Gln Ser Ile Leu Asp
199 65 70 75 80
200 Glu Phe Leu Gln Thr Tyr Gly Ser Leu Ile Pro Leu Ser Thr Asp Glu
201 85 90 95
202 Val Val Glu Lys Leu Glu Asp Ile Phe Gln Gln Glu Phe Ser Thr Pro
203 100 105 110
204 Ser Arg Lys Gly Leu Val Leu Gln Leu Ile Gln Ser Tyr Gln Arg Met
205 115 120 125
206 Pro Gly Asn Ala Met Val Arg Gly Phe Arg Val Ala Tyr Lys Arg His
207 130 135 140
208 Val Leu Thr Met Asp Asp Leu Gly Thr Leu Tyr Gly Gln Asn Trp Leu
209 145 150 155 160
210 Asn Asp Gln Val Met Asn Met Tyr Gly Asp Leu Val Met Asp Thr Val
211 165 170 175
212 Pro Glu Lys Val His Phe Phe Asn Ser Phe Phe Tyr Asp Lys Leu Arg
213 180 185 190
214 Thr Lys Gly Tyr Asp Gly Val Lys Arg Trp Thr Lys Asn Val Asp Ile
215 195 200 205
216 Phe Asn Lys Glu Leu Leu Leu Ile Pro Ile His Leu Glu Val His Trp
217 210 215 220
218 Ser Leu Ile Ser Val Asp Val Arg Arg Arg Thr Ile Thr Tyr Phe Asp
219 225 230 235 240
220 Ser Gln Arg Thr Leu Asn Arg Arg Cys Pro Lys His Ile Ala Lys Tyr
221 245 250 255
222 Leu Gln Ala Glu Ala Val Lys Lys Asp Arg Leu Asp Phe His Gln Gly

```

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/848,852A

TIME: 11:52:09

Input Set : N:\Crf3\RULE60\09848852.raw

Output Set: N:\CRF3\12202001\I848852A.raw

```

223           260           265           270
224 Trp Lys Gly Tyr Phe Lys Met Asn Val Ala Arg Gln Asn Asn Asp Ser
225           275           280           285
226 Asp Cys Gly Ala Phe Val Leu Gln Tyr Cys Lys His Leu Ala Leu Ser
227           290           295           300
228 Gln Pro Phe Ser Phe Thr Gln Gln Asp Met Pro Lys Leu Arg Arg Gln
229           305           310           315           320
230 Ile Tyr Lys Glu Leu Cys His Cys Lys Leu Thr Val
231           325           330

```

233 (2) INFORMATION FOR SEQ ID NO: 4:

235 (i) SEQUENCE CHARACTERISTICS:

236 (A) LENGTH: 1991 base pairs

237 (B) TYPE: nucleic acid

238 (C) STRANDEDNESS: single

239 (D) TOPOLOGY: linear

241 (vii) IMMEDIATE SOURCE:

242 (A) LIBRARY: BEPINOT01

243 (B) CLONE: 2056178

245 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

247 GCCTCCCTGT CCCCCGACCC TCTTTTGATG CCTCAGCAAG TGAAGAGGAG GAAGAAGAGG 60
248 AGGAGGAGGA GGATGAAGAT GAAGAGGAGG AAGTGGCAGC TTGGAGGCTG CCCCCAAGAT 120
249 GGAGTCAGCT GGGAACTCC CAGCGGCCCC GCCCTTCCCG CCCCCTCAT CGAAAAACCT 180
250 GCTCACAGCG CCGCCGCCGA GCCATGAGAG CCTTCCGGAT GCTGCTCTAC TCAAAAAGCA 240
251 CCTCGCTGAC ATTCCACTGG AAGCTTTGGG GGCGCCACCG GGGCCGGCGG CGGGGCCTCG 300
252 CACACCCCAA GAACCATCTT TCACCCCAGC AAGGGGGTGC GACGCCACAG GTGCCATCCC 360
253 CCTGTTGTCG TTTTGACTCC CCCCAGGGGC CACCTCCACC CCGGCTGGGT CTGCTAGGTG 420
254 CTCTCATGGC TGAGGATGGG GTGAGAGGGT CTCCACCAGT GCCCTCTGGG CCCCCCATGG 480
255 AGGAAGATGG ACTCAGGTGG ACTCCAAAGT CTCCTCTGGA CCCTGACTCG GGCCTCCTTT 540
256 CATGTACTCT GCCCAACGGT TTTGGGGGCA AATCTGGGCC AGAAGGGGAG CGCCTTGGC 600
257 ACCCCCTGAT GCCAGCATCT TCATCAGCAA TGTGTGCAGC ATCGGGGACC ATGTGGCCCA 660
258 GGAGCTTTT CAGGGCTCAG ATTTGGGCAT GGCAGAAGAG GCAGAGAGGC CTGGGGAGAA 720
259 AGCCGGCCAG CACAGCCCCC TGCGAGAGGA GCATGTGACC TGCGTACAGA GCATCTTGGA 780
260 CGAATTCCTT CAAACGTATG GCAGCCTCAT ACCCCTCAGC ACTGATGAGG TAGTAGAGAA 840
261 GCTGGAGGAC ATTTTCACAG AGGAGTTTTC CACCCCTTCC AGGAAGGGCC TGGTGTGCA 900
262 GCTGATCCAG TCTTACCAGC GGATGCCAGG CAATGCCATG GTGAGGGGCT TCCGAGTGGC 960
263 TTATAAGCGG CACGTGCTGA CCATGGATGA CTTGGGGACC TTGTATGGAC AGAACTGGCT 1020
264 CAATGACCAG GTGATGAACA TGTATGGAGA CCTGGTCATG GACACAGTCC CTGAAAAGGT 1080
265 GCATTTCTTC AATAGTTTCT TCTATGATAA ACTCCGTACC AAGGGTTATG ATGGGGTGAA 1140
266 AAGGTGACC AAAAACGTGG ACATCTTCAA TAAGGAGCTA CTGCTAATCC CCATCCACCT 1200
267 GGAGGTGCAT TGGTCCCTCA TCTCTGTTGA TGTGAGGCGA CGCACCATCA CCTATTTTGA 1260
268 CTCGAGCGT ACCCTAAACC GCCGCTGCC TAAGCATATT GCCAAGTATC TACAGGCAGA 1320
269 GGCGGTAAAG AAAGACCGAC TGGATTTCCA CCAGGGCTGG AAAGGTTACT TCAAAATGAA 1380
270 TGTGGCCAGG CAGAATAATG ACAGTGAAGT TGGTGCTTTT GTGTTGCAGT ACTGCAAGCA 1440
271 TCTGGCCCTG TCTCAGCCAT TCAGCTTCAC CCAGCAGGAC ATGCCCAAAC TTCGTCGGCA 1500
272 GATCTACAA GAGCTGTGTC ACTGCAAAC CACTGTGTGA GCCTCGTACC CCAGACCCCA 1560
273 AGCCCAATAA TGGGAAGGGA GACATGGGAG TCCCTTCCCA AGAACTCCA GTTCCTTTCC 1620
274 TCTCTGCTT CTTCCCACTC ACTTCCCTTT TGTCTTTTCA ATTTAAATGT TTCAATTCT 1680
275 GTATTTTTTT TCTTTTGAGA GAATACTTGT TGATTTCTGA TGTGCAGGGG GTGGCTACAG 1740
276 AAAAGCCCCT TTCTTCCTCT GTTTGCAGGG GAGTGTGGCC CTGTGGCCTG GGTGGAGCAG 1800

```

VERIFICATION SUMMARY

DATE: 12/20/2001

PATENT APPLICATION: US/09/848,852A

TIME: 11:52:10

Input Set : N:\Crf3\RULE60\09848852.raw

Output Set: N:\CRF3\12202001\I848852A.raw

L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]